

Human Uniqueness, Cognition by Description, and Procedural Memory

John Bolender, Burak Erdeniz & Cemil Kerimoğlu

Evidence will be reviewed suggesting a fairly direct link between the human ability to think about entities which one has never perceived — here called ‘cognition by description’ — and procedural memory. Cognition by description is a uniquely hominid trait which makes religion, science, and history possible. It is hypothesized that cognition by description (in the manner of Bertrand Russell’s ‘knowledge by description’) requires variable binding, which in turn utilizes quantifier raising. Quantifier raising plausibly depends upon the computational core of language, specifically the element of it which Noam Chomsky calls ‘internal Merge’. Internal Merge produces hierarchical structures by means of a memory of derivational steps, a process plausibly involving procedural memory. The hypothesis is testable, predicting that procedural memory deficits will be accompanied by impairments in cognition by description. We also discuss neural mechanisms plausibly underlying procedural memory and also, by our hypothesis, cognition by description.

Keywords: basal ganglia; cerebellum; cognition by description; knowledge by description; language evolution; procedural memory; theory of descriptions

1. Introduction

We review evidence suggesting a fairly direct link between the human ability to think about entities which one has never perceived, what we call ‘cognition by description’,¹ and procedural memory. The discussion is exploratory and its conclusions are tentative, but we submit that evident links between specific forms of uniquely human cognition and procedural memory merit attention.

For helpful comments on an earlier draft and/or email exchanges on pertinent topics, we would like to thank Noam Chomsky, Ayşe Elif Erson, Daniel Everett, Andrew Nevins, Steven Pinker, Üner Tan, Michael Ullman, and two anonymous referees for this journal. Any remaining deficiencies are due to the authors alone.

¹ We use the term ‘cognition by description’ — instead of Russell’s (1910) term ‘knowledge by description’ — because knowledge is often taken to involve justification and hence to have normative implications (Kim 1993: chap. 12). Since our aim is scientific, we choose a purely descriptive term.



2. Cognition by Description and Merge

Non-humans give no evidence of being able to think about fictional characters or entities which they have never perceived, such as electrons or Poseidon or the children of one's unborn children. Instead, one finds a one-to-one mapping between a brain process and some property in the environment (Gallistel 1990). In spoken remarks (Chomsky 2006), Noam Chomsky has reflected on how this differs from the semantics of human mental systems:

The concepts and words that [humans] have do not pick out entities in the world that, say, a physicist could identify, that can be found by mind-external investigation. But, rather, they are basically products of the imagination. [...] The formal concept of reference that you study in logic, that's developed by Frege, Peirce, Tarski, Carnap, and so on — it just doesn't apply to natural language. [...] Human language and thought just don't have [...] terms that pick out pieces of the world. This appears to be a respect in which humans differ sharply from any other organism. As far as is known, animal communication systems are based on [...] an isomorphism between some internal symbol and some identifiable aspect of the world. [...] Looking at the vervet monkey anthropomorphically, the way we look at everything, we say that the monkey is giving this call because there's an eagle there, and it's telling the rest of the crew of monkeys to run away. What appears to be happening [however] is that there's a reflexive reaction to some, say, motion of leaves or something like that, and then the call comes out. There's apparently an identifiable isomorphic relation between the call and a physically identifiable aspect of the environment. [...] That appears to be the case for all animal communication systems, as far as anyone's discovered. If that's the case, then one fundamental difference between humans and the rest of the organic world is that we have concepts that do not pick out mind-independent entities.

How can humans use symbols, ultimately mental symbols, to indicate things in one's 'subjective universe', as it is sometimes called?

This ability should not be equated with displaced reference. A typical definition of displaced reference would be "the ability to refer to information that is spatially and temporally displaced from the location of the speaker and the listener" (Morford & Goldin-Meadow 1997: 420). But this definition doesn't distinguish referring to entities that one perceived earlier, but which are not currently present, from the designation of entities that one has *never* perceived. There is an important difference between seeing a red ball, leaving the room so that one no longer sees it, and then referring to the red ball; as opposed to referring to the children of one's unborn children. There is also the ability to talk and think about entities which do not even exist. Strictly speaking, this is not *reference* at all, but it is a sophisticated semantic ability nonetheless.

In this article we are concerned with what appears to be a uniquely human ability, namely the ability to talk and think about entities which one has never perceived and which, in some cases, do not exist. This is what we call 'cognition by description', a naturalized variant of Bertrand Russell's knowledge by description (Russell 1910, 1959). Chomsky has suggested that a computational operation could explain this uniquely human ability (Hauser *et al.* 2002, Chomsky

2005, 2006, 2007). In other words, the evolution of a recursive procedure, in conjunction with whatever mental apparatus was already there,² resulted in cognition by description.

Derek Bickerton, however, insists that the mere addition of recursion in hominid evolution is not enough to explain this sophisticated capacity.

Those [semantic] properties, as [Chomsky] quite correctly states, are precisely the properties that distinguish human concepts from the concepts of other species — they refer to mental constructs rather than natural objects. But if concepts with such properties are unique to human language, *how could recursion have applied to them when language did not yet exist?* Either those concepts (and probably the words with which they were linked) already existed, implying some kind of system intermediate between animal communication and true language, or recursion could not [...] have applied to anything. Since syntactic language now exists, it is a logically unavoidable conclusion that there must have been some kind of protolanguage before recursion.
(Bickerton 2005: 3, emphasis in the original)

Not surprisingly, Bickerton takes the following two issues in language evolution to be fundamentally distinct:

- (1) How did symbolic units (words or manual signs) evolve?
- (2) How did syntax evolve?

Symbolic units and syntax are the only real novelties in human communication, and are therefore the most salient (as well as the most difficult) of the things any adequate theory of language evolution must account for. There is no reason to believe that the emergence of the two was either simultaneous or due to similar causes, and some good reasons for supposing the contrary. Chomsky (1980) has made a clear distinction between the conceptual and the computational aspects of language. The conceptual element, especially if we take it to include conceptual structure as well as the lexical instantiation of the latter, must be phylogenetically far older than any computational mechanism.
(Bickerton 2007: 511)³

Contra Bickerton, we wish to show how the addition of a specific recursive operation could transform a system of mental symbols which are referential in the manner of, say, vervet communication systems, into a symbol system suitable for cognition by description. It is not necessary to posit an a-grammatical protolanguage, the concepts of which made possible cognition by description, prior to the evolution of uniquely human computational abilities. In other words, Bickerton's (1) and (2) may be far more tightly intertwined than he recognizes, where the "symbols" in his (1) designate entities in one's 'subjective universe'. Contrary to Bickerton, adding the right sort of computations to a purely referen-

² "In conjunction with whatever mental apparatus was already there" is an important qualification. In discussing this topic, people sometimes forget that it is a *necessary* condition that is at issue, not a sufficient one. This is why a condition, such as Williams syndrome (Karmiloff-Smith 1992, Smith 2004), in which mental retardation is accompanied by sophisticated syntax is not a counterexample to the hypothesis of this paper.

³ One can recognize the distinction between conceptual and computational systems, while acknowledging that one system uses symbols generated by the other.

tialist system can yield a mind capable of cognition by description.⁴ All the really heavy lifting here, in fact, was done by Russell (1905, 1919) in his theory of descriptions and his closely related work on knowledge by description (Russell 1910, 1959), and also by Chomsky (1976) in his work on trace theory. We try to show here how Russell's insights can be extended to questions of human cognitive evolution.⁵

In doing so, we assume that the computational core of language plays a large role in uniquely human cognition generally, a simpler hypothesis than positing one recursive system for language and a separate one for belief-forming systems. This working hypothesis suggests a working methodology, namely to investigate syntactic computations as a means of understanding uniquely human concepts. It wouldn't hurt to emphasize this working methodology, and to write it out as a principle:

So far as possible, seek explanations of uniquely human concepts in terms of syntactic computations.

A classic example of the application of this methodology would be Chomsky's attempt to understand unbounded counting in terms of syntax (Hauser *et al.* 2002, Chomsky 2005), to which we return later.

Russell's theory of descriptions is often understood as a theory about the logical form of sentences containing determiner phrases (Neale 1990).⁶ Logical form is here defined as "whatever features of sentence structure (1) enter directly into semantic interpretation of sentences and (2) are strictly determined by properties of (sentence-) grammar" (Chomsky 1976: 305-306).⁷

Russell was especially concerned with the logical form of sentences containing determiner phrases which designate at most one thing; for example, such phrases as *the inventor of the telegraph*, *the author of De Legibus*, *Whistler's mother*, and *my favorite book*. This class of determiner phrases also includes phrases which designate at most one type of thing, such as *the element with atomic number 1*. These phrases are known as 'definite descriptions'. Russell (1905, 1919) was concerned to show that sentences containing definite descriptions, which we

⁴ This would imply that humans and non-human primates both utilize symbols, 'primitives', with referentialist semantics. This disagrees with Chomsky's (2007: 20) remark that "even the simplest words and concepts of human language and thought lack the relation to mind-independent entities that has been reported for animal communication". But it is not clear how Chomsky means to defend such a sweeping claim. Our proposal is that, while humans and other primates share a referentialist semantic system, humans alone enjoy a recursively generated semantics which draws its primitives from that referentialist system.

⁵ In doing so, we take no stand on the metaphysical issues which concerned Russell, such as sense-datum theory, logical data, etc.

⁶ An alternative would be to interpret Russell's theory as concerned with an ideal or perfect language, as opposed to natural language. Russell was not always consistent on this point, but he sometimes explicitly denied any interest in natural language: "My theory of descriptions was never intended as an analysis of the state of mind of those who utter sentences containing descriptions" (Russell 1957: 388). But if the theory of descriptions can be given a testable formulation and is shown to have explanatory power, e.g., in accounting for human uniqueness, than it deserves to be taken seriously as a scientific hypothesis about the human mind.

⁷ This is Chomsky's definition of 'LF'. How precisely Russell's own notion of logical form overlaps and contrasts with LF will not be discussed.

will call 'definites', implicitly feature bound variables. Using a simplified example, the logical form of the definite *The author of De Legibus was Roman* would be [*There is a unique x , such that x wrote De Legibus*] and [*For all x , if x wrote De Legibus then x was Roman*]. The logical form more abstractly stated would be [*There is a unique x , such that Lx*] and [*for all x if Lx then Rx*], where L and R are predicates. The semantic interpretation could be expressed as *Whatever unique thing wrote De Legibus, that thing was Roman*.

According to Russell's (1910, 1959) theory of knowledge by description, the ability to have thoughts of the form [*There is a unique x , such that Lx*] and [*for all x if Lx then Rx*] makes it possible to think about things one has never perceived. For example, one can think about the author of *De Legibus*, namely Cicero, by forming quantified mental representations even if one had never met him. This is because the use of quantification can restrict the extension of a predicate to at most one individual thus forming a predicate which can do much of the semantic work that would otherwise be done by a proper name. Russell was, at least implicitly, following our working methodology, since he submitted that the logical form of definites also enters into knowledge by description. In other words, the operator-variable structures which figure into language also figure in the formation of thoughts about unobserved things. Analogously, we suggest that bound variables play a crucial role in cognition by description, such as one's beliefs about Cicero. This is simpler than supposing that the mind reinvents the wheel, so to speak, by generating operator-variable structures for language and then generating them separately, and all over again, for the belief systems.

To illustrate, suppose that an early hominid discovers an artifact in the forest, say, a stone tool. Let's suppose that this hominid uses the sort of mental symbol system that plausibly characterizes vervet monkeys, so that the system's semantics is limited to objects of immediate experience. The hominid thinks about the tool only by reason of "a reflexive reaction", to use Chomsky's phrase. In other words, there is a fairly direct causal link between the presence of the tool and the tokening of the relevant mental symbol. Now let us also suppose that the maker of the tool died long before its discoverer was born, and no bodily remains of that maker are anywhere to be found nearby. Given a system of mental representation limited to currently existing entities — essentially a reflex — the discoverer might be able to think about the tool but not about its maker.

But we know that that's not what actually happens, at least not for *Homo sapiens*. The *Homo sapiens* can think about the maker of the tool without needing to perceive that maker. How is this possible? Russell suggested that one uses definite descriptions. Since a definite description is implicitly a quantifier phrase, the quantifier operator restricts the extension of the relevant predicate so that it is satisfied by at most one individual or set of individuals singled out by their shared properties, the resulting expression being equivalent, for practical purposes, to a symbol for an unperceived entity or type of thing. To return to the example of the hominid encountering the stone tool; s/he could designate the long dead maker of the tool by forming a mental representation with the logical form *there is a unique x , such that x made this tool*.

Note that Russell was suggesting that one uses the product of language, namely a definite description, in mental operations that one would not intuitively

think of as linguistic. To use the terminology of modularity theory (Fodor 1983), one uses representations generated by the language faculty to form representations in the belief systems. This explanation is more economical than proposing two distinct systems for generating operator–variable structures, and thus agrees with our working methodology.⁸

From a Russellian perspective, adding variable binding to an otherwise referentialist symbol system would suffice for definite descriptions and hence knowledge by description. Therefore, *if the addition of a recursive procedure can make possible variable binding*, then one would have to disagree with Bickerton when he says that one cannot account for the evolution of cognition by description by appealing to the evolution of a recursive procedure. So can the mere addition of a recursive procedure explain variable binding? To answer the question, let's consider what sort of recursive procedure Chomsky has in mind in the first place.

In this procedure, known as *Merge*, two objects are combined such that one alone is the 'head', determining the resultant object's combinatorial properties (Chomsky 2005, 2007). The resulting compound can then be merged with another object to yield a more complex compound with a new head. And so on. This makes possible the recursive embedding of an object within an object of the same type, such as a verb phrase within a verb phrase.

For example, *for* can be merged with *mercy* to yield a PP:

- (1) [PP *for mercy*]

For is the head, since it determines that the phrase is a Preposition Phrase, rather than being a Noun Phrase. The grammatical category of the phrase determines how it can combine with other objects. If it were a Noun Phrase, it would appear in different grammatical contexts. Note that this result of Merge can itself be merged with *plead* to form the Verb Phrase:

- (2) [VP *plead* [PP *for mercy*]]

This VP can be merged with *to* thus forming the Infinitive Phrase:

- (3) [IP *to* [VP *plead* [PP *for mercy*]]]

This IP can further be merged to the verb *refuse* yielding the next VP:

- (4) [VP *refuse* [IP *to* [VP *plead* [PP *for mercy*]]]]

And so on.

⁸ An anonymous referee suggested that our hypothesis is committed to there being an implausible isomorphism between language and the systems of belief, specifically that language produces belief structures and that the belief systems produce syntactic structures. To the contrary, we posit a division of labor and hence a crucial difference between these two faculties. One would be crippled without the other precisely because they are not isomorphic.

We see here an illustration of recursion, in this case the inclusion of a Verb Phrase in a Verb Phrase. The result of this repeated use of Merge is a hierarchical structure in which one phrase dominates another which dominates another, in this case terminating with the domination of *for* and *mercy*.⁹ There is evidence that language has hierarchical phrase structure (Miller 1962, Miller & Isard 1963, Epstein 1961a, 1961b, Fodor & Bever 1965, Johnson 1965, Graf & Torrey 1966, Mehler *et al.* 1967, Anglin & Miller 1968, Levelt 1970). Furthermore, positing a recursive operation is unavoidable in explaining how finite resources can potentially generate an infinite number of hierarchically ordered structures composed of discrete elements (Turing 1950, Boolos *et al.* 2002).

Merge can only take two forms: Either an object *O* is merged to an object which is a constituent of *O* or *O* is merged to an object which is not a constituent of *O* — internal Merge and external Merge, respectively (Chomsky 2005). The earlier example of Merge, illustrated in *refuse to plead for mercy*, was external. What about internal Merge? Consider the phrase *Socrates thought what*. Merging *what* with *Socrates thought what* to yield *what Socrates thought what* is an example of internal Merge, because *what* was already a constituent of *Socrates thought what*. The resulting four-part phrase *what Socrates thought what* would not be fully pronounced. But semantically, all four elements are interpreted, the first *what* as an operator and the second *what* as the variable it binds.¹⁰ In other words, the semantic interpretation would be *for which thing x, Socrates thought x*, pronounced in English as ‘What Socrates thought’.

[A]s a simple matter of logic, there are two kinds of Merge, external and internal. External Merge takes two objects, say *eat* and *apples*, and forms the new object that corresponds to *eat apples*. Internal Merge — often called Move — is the same, except that one of the objects is internal to the other. So applying internal Merge to *John ate what*, we form the new object corresponding to *what John ate what, ...*. [A]t the semantic interface, both occurrences of *what* are interpreted: The first occurrence as an operator and the second as the variable over which it ranges, so that the expression means something like *for which thing x, John ate the thing x*. At the sensorimotor side, only one of the two identical syntactic objects is pronounced, typically the structurally most salient occurrence. (Chomsky 2007: 21)

Internal Merge generally creates operator–variable structures (Chomsky 1976).

Chomsky (2005) argues for Merge by noting that the arrangement of discrete elements into a potential infinitude of hierarchical structures requires a combinatorial operation. Given that Merge as such is a recursive operation and that internal Merge accounts for bound variables, one can begin to see how recursion could figure into ‘definites’ (i.e. sentences containing definite descriptions) and hence cognition by description, as per our working methodology.

More needs to be said as to exactly how internal Merge figures into

⁹ This process is indifferent between synthesis and analysis. Syntax merely distinguishes grammatical from ungrammatical structures, it is not specifically a sentence producing process. For some early (pre-Merge) discussion of this, see Chomsky (2002: 48).

¹⁰ By copying *what* into a superordinate position, it binds the unspoken *what* in the subordinate position, according to the c-command condition on binding (Reinhart 1976).

definites, but before doing so, we need to reflect further on the nature of definites. Here we follow Larson & Segal's (1998: 247f.) analysis of quantified sentences including definites. A quantified sentence is analyzable into three elements:

- (5) a. A **quantification** stating how many are involved, e.g., *all*. In the case of a definite, the quantifier states that at most one is involved, e.g., *the*.
- b. A **restriction** stating the class of entities involved, e.g., *mice*.
- c. A **scope** stating what is true of the individuals in the restriction, e.g., *being mortal*.

In *All mice are mortal*, *all* expresses quantification, *mice* expresses the restriction, and *are mortal* expresses the scope. It will be seen that internal Merge plays a role in distinguishing restriction from scope in sentences containing determiner phrases, including definites.

Many linguists today agree with Russell that the use of definites involves operator–variable structures at the level of logical form,¹¹ although the Russellian approach to such structures has been amended somewhat (Keenan & Stavi 1986, Neale 1990, Larson & Segal 1998). While Russell used the unary quantifiers introduced by Frege, it is more plausible that determiners in natural languages are binary; that *all*, for example, is a relational predicate whose arguments are two sets. For example, the *all* in *All mice are mortal* expresses a relation between the set of mice and the set of mortals. But this approach to determiners is still Russellian in an important respect; namely, it still involves bound variables. Furthermore, the bound variables that figure into sentences using determiners are plausibly due to a sub-case of internal Merge known as 'quantifier raising' (see Larson & Segal 1998, Hornstein & Uriagereka 1999 for recent discussion). Given our working methodology, this means that it is a plausible hypothesis that quantifier raising plays a crucial role in cognition by description.

According to recent semantic theory (*supra*), two different sorts of Merge must occur so as to distinguish restriction from scope. Consider the logical form of *All mice are mortal* and how it is produced. *All* is externally Merged to *mice* so that *mice* serves as the restriction. But an internal Merge operation must be performed so that *being mortal* will serve as scope. Internal Merge establishes relations of scope, in the case of quantification, by producing bound variables. More specifically, *all mice* initially occurs as the complement of *mortal* and is then internally merged in a higher position (i.e. 'raised') leaving an unpronounced variable as the complement of *mortal*, namely *mortal* x. The result of this quanti-

¹¹ Donnellan (1966, 1968) argues that there is a 'referential use' of definite descriptions such that a definite need not contain a bound variable (see also Devitt 2004). This is often taken as "an attack on Russell", but Donnellan is only saying that Russell's theory of descriptions doesn't fully generalize. Donnellan never denied that definites sometimes conceal operator-variable structures as revealed on a semantic analysis. This would be the case for the 'attributive use' of definites. So, even if Donnellan is right, it does not mean that the semantic analyses of this paper are wrong. It would just mean that they are only true of a specific use of definites. That 'attributive use' would employ the same computational procedures (e.g., internal Merge) which enter into cognition by description.

fier raising is that the set of mortal things becomes the scope of the expression. Using Neale's (1990) notation, '[all x : mice x] (mortal x)' is a good expression of the logical form of *All mice are mortal*; 'all x ' must be raised to a superordinate position in order to bind the variable in 'mortal x '.¹² Let's consider an example of a definite, namely, *The maker of this arrowhead was skilled*. Ignoring tense, the logical form is more revealing notated as follows: [the x : make this arrowhead x] (skilled x). The variable appearing in '(skilled x)' is bound by the operator as a result of quantifier raising, the set of skilled things thus serving as the scope.

If internal Merge is necessary for cognition by description does it follow that internal Merge is also necessary for formulating counterfactuals or questions at the level of thought? Not necessarily. In the case of counterfactuals and questions, all that is needed is the ability to combine meaningful units in various different ways. Assume that Socrates, Thrasymachus, and kissing are objects of immediate awareness. One should be able to form the thought *Socrates kissed Thrasymachus*, even if it is a false thought, simply by combining the relevant meaningful units so as to yield the representation *Socrates kissed Thrasymachus*. This would not be cognition by description, but it would be counterfactual.¹³ One should also be able to formulate the query *Did Socrates kiss Thrasymachus?* by forming the representation *Socrates kissed Thrasymachus* and then adding the conceptual element of interrogation.¹⁴

As briefly noted earlier, there is a precedent for attempting to explain uniquely human cognitive abilities in syntactic terms in Chomsky's (2005) suggestion that unbounded counting results from Merge. As Noam Chomsky (p.c.) puts it, "[t]here are a number of ways of deriving the number system from Merge. To take one, assume that the lexicon has a single member, call it 1, and accept the convention that $\{X\} = X$. Then $1 = \{1\}$. Internal Merge yields $\{1, \{1\}\}$. Call it 2. Etc. Addition and other operations follow pretty simply". This agrees with our working methodology.

3. Unique to Humans?

In the past few years, there has been much discussion as to whether recursion in cognitive processes is unique to humans (Hauser *et al.* 2002, Pinker & Jackendoff 2005, Parker 2006) or whether a specific recursive procedure, such as Merge, is unique to humans (Chomsky 2005). The debate is relevant here because we want to know which is the more plausible hypothesis: Did the evolution of Merge as such usher in cognition by description, or was it specifically the evolution of internal Merge? Or maybe even just quantifier raising? If Merge-like procedures are found in other species, but without evidence of internal Merge, this would be

¹² Raising is necessary for binding because of the c-command condition.

¹³ Note that Russell's knowledge by description, even though it involves knowledge of some truths, still counts as knowledge of things. For Russell (1959: 46f.), I have knowledge by description of, say, Socrates and Thrasymachus, but I do not have knowledge by description, say, that Socrates pitied Thrasymachus.

¹⁴ This may not be the same as forming the sentence *Did Socrates kiss Thrasymachus?* which evidently does require internal Merge, at least in English, with the unspoken trace of *did* following *Socrates*.

relevant. The debate concerning whether or not recursion is unique to humans, and the closely related question of whether or not hierarchically structured mental representations are unique to humans, remain very much alive (Gibson 1993, Byrne & Russon 1998, Spinozzi *et al.* 1999, Bergman *et al.* 2003, McGonigle *et al.* 2003, Fitch & Hauser 2004, Suzuki *et al.* 2006).

Here is one example of the debate: Some scientists have argued that the European starling can parse recursively center-embedded structures. Starlings can be trained to behave as though they have internalized rules of the form $a^n b^n$ as applied to the chirps and warbles they are familiar with from their own songs, at least when $n=2$ (Gentner *et al.* 2006). Does this mean that they are parsing structures of the form $[A[AB]B]$ in which there is an AB recursively nested in another token of AB ? In other words, does it mean that we find here the computational power minimally required for a context-free grammar as Timothy Gentner concludes? Not necessarily. According to Chomsky, the conclusion of Gentner and his colleagues “is based on an elementary mathematical error” (quoted in Goudarzi 2006). He adds that the birds’ behavior “has nothing remotely to do with language; probably just with short-term memory”. In other words, the starlings could be employing a non-recursive device for counting chirps and warbles. The bird could be counting two chirps, storing the result in memory, and then checking to see if the warbles also equal two.¹⁵ This need only bestow on them the computational power of a finite-state automaton with counters (Chomsky 1959: 151), a non-recursive machine.

We do not take a position as to whether recursion is unique to humans. But we do hypothesize that (at least) internal Merge is unique to humans, and that this explains why cognition by description is only found among them. In fact, the limitation of cognition by description to humans is evidence for the limitation of internal Merge to humans. (When we say ‘humans’, we are not excluding other extinct hominid species; we take no stand on whether, say, Neanderthals utilized internal Merge.)

Hypothesizing that internal Merge is unique to humans leaves open the question of whether or not external Merge is as well. Fitch *et al.* (2005: 186-187) have considered the possibility that navigation in some nonhuman species employs a combinatorial computational procedure which is very much like external Merge. To give an example, an animal may be able to remember the location of its home by means of a mental representation that would be well expressed in English as $[[[the\ hole]\ in\ the\ ground]\ near\ the\ tree]\ by\ the\ lake]$, exhibiting a nested structure analogous to $[refuse\ [to\ [plead\ [for\ mercy]]]]$ and also exhibiting compositionality, an important feature of Merge. But note that internal Merge is not required to form this specific mental representation. There is a tendency for linguists working in the minimalist paradigm to treat internal and external forms of Merge as necessarily both being utilizable by a mind if either is (Berwick 1998). But, given that internal Merge requires a more developed procedural memory system than does external Merge alone, as we will discuss in the next section, it should come as no surprise that a mind may utilize the external form only.

¹⁵ See Pinker & Jackendoff (2005, fn. 10) for a similar criticism, albeit directed against Fitch & Hauser (2004).

4. Internal Merge and Types of Memory

Internal Merge is tantamount to what is often called ‘syntactic movement’ or just ‘movement’. This is because internal Merge does look like the rearrangement of parts, if one focuses on phonology alone. For example, in the case of internally merging *what* with *Socrates thought what*, it looks as though *what* moves from the end of the structure to its beginning, a transformation of a more basic structure. Syntactic movement respects parts of speech and phrase structure; i.e. it is ‘structure-dependent’ meaning that part of speech and phrasal location are crucial in determining which object is moved. In *The dog who dug there was growling*, it is possible to move *was* to the front yielding *Was the dog who dug there growling?*. But *dug* cannot be moved to the front. So **Dug the dog who there was growling?*, despite its lovely poetic meter, is ungrammatical. Not only is it the case that one can only move a verb in English question formation, it also matters which clause the verb appears in prior to movement. It is the auxiliary verb in the main clause which moves. Given poverty-of-the-stimulus evidence collected by Stromswold (1999), the structure dependency of movement seems to be innate, and hence an invariant feature of language.

To know how a sentence is divided into phrases, and the parts of speech of its elements, is to remember something about how it was constructed, i.e. how objects were merged together to form this complex object, this sentence. To know, for example, that *The dog who dug there was growling* contains a sub-clause, and where that sub-clause begins and ends, is to remember that *The dog who dug there was growling* was put together out of simpler parts and to remember something about what the parts of speech of those parts were at each step in the derivation. Also to know that *was* is here an auxiliary verb is to know something about how the parts of the sentence were put together. A mapping from one derivational step to the next, when movement is involved, “rearranges the elements of the string to which it applies, and it requires considerable information about the constituent structure of this string” (Chomsky 1956: 121). This information is tantamount to a memory of derivational steps, what is sometimes called ‘derivational memory’. In other words, internal Merge requires derivational memory (Piattelli-Palmarini & Uriagereka 2005: 53f.). *If cognition by description involves quantifier raising and quantifier raising is a sub-case of internal Merge, then cognition by description requires derivational memory.*

We need to reflect on some more general features of memory before returning to the discussion of the specific memory demands of internal Merge. When the word ‘memory’ is used in everyday language, it is usually declarative memory that is meant; i.e. the conscious recollection of facts and events. There are also unconscious, evidently, non-declarative memory systems too (Squire 2004). The form of non-declarative memory of special interest in understanding the structure-dependency of internal Merge is procedural memory, namely the sort of memory implicated

in the learning of the new, and the control of long-established, motor and cognitive ‘skills’, ‘habits’, and other procedures, such as typing, riding a bicycle, and skilled game playing [...]. The [procedural] system underlies aspects of rule-learning [...], and is particularly important for acquiring and

performing skills involving sequences — whether the sequences are serial or abstract, or sensori-motor or cognitive [...]. It is commonly referred to as an implicit memory system because both the learning of the procedural memories and the memories themselves are generally not available to conscious access.

(Ullman & Pierpont 2005: 401; emphasis added — JB, BE & CK)

How might this relate to language? Linguistic mappings of sounds onto meanings can be divided into the idiosyncratic and the principled. For example, *refuse to plead for mercy* is mapped onto its semantic content in a principled way because its meaning is a function of the meanings of its parts and their manner of combination. That's compositionality. The same cannot be said for *kick the bucket* when used as an idiom. One must memorize the meaning of the latter, rather than constructing it from its parts.¹⁶ Michael Ullman and his colleagues hypothesize that principled mappings utilize the procedural memory system, while idiosyncratic mappings utilize declarative memory, what is known as the 'declarative/procedure model' (Ullman & Gopnik 1994, Pinker & Ullman 2002, Ullman 2004, Ullman & Pierpont 2005, Newman *et al.* 2007). In terms of Chomsky's linguistics, this would mean that Merge requires procedural memory, whereas lexical pairings of sound and meaning utilize declarative memory.

Part of what recommends Ullman's hypothesis is its accounting for the otherwise mysterious range of disabilities associated with Specific Language Impairment (SLI), "a developmental disorder of language in the absence of frank neurological damage, hearing deficits, severe environmental deprivation, or mental retardation" (Ullman & Pierpont 2005: 399). The authors note that, in addition to difficulties in grammar, those with SLI exhibit impairments in motor skills, working memory, and word retrieval. This cluster of symptoms could be explained in terms of a deficit in procedural memory. The hypothesis is further recommended by the fact that disorders involving impairment of procedural memory are accompanied by grammatical difficulties, while disorders involving declarative memory are accompanied by lexical difficulties (Ullman 2004). It has also been hypothesized that a role for the *FOXP2* gene in procedural memory may explain why a defect in that gene results in grammatical difficulties (Ullman & Pierpont 2005, Piattelli-Palmarini & Uriagereka 2005), a basal ganglia abnormality being implicated (Watkins *et al.* 2002).

The procedural/declarative hypothesis is controversial. Some have argued that the evidence favors a single-mechanism model (Bird *et al.* 2003, Joanisse & Seidenberg 1999, McClelland & Patterson 2002, Longworth *et al.* 2005). Newman *et al.* (2007: 436) conclude that "the issue is still open, and further evidence is necessary to help constrain the range of possible theoretical interpretations". The procedural/declarative model is assumed here for the sake of developing a hypothesis to test.

The structure of a phrase involves an abstract sequencing insofar as it exhibits hierarchical relations, as illustrated earlier by the example *refuse to plead for mercy*. So, on Ullman's model, Merge requires procedural memory. This point

¹⁶ Although it is principled to the extent that one can say *kicks the bucket* or *kicked the bucket*, and so on.

is essentially the same as that made by Ullman & Pierpont (2005) in their discussion of 'rule governed' syntax. But note that Piattelli-Palmarini & Uriagereka (2005) point out that derivational memory, discussed above, is also plausibly procedural since it is a kind of sequence memory. Memory of derivational steps is a memory of the order in which objects were merged together and which parts of speech those objects were. *So internal Merge should place even greater demands on procedural memory than external Merge alone.* External Merge alone involves hierarchical relations, but internal Merge also requires a memory of the steps taken in forming such relations (Chomsky 2002: 37). The role of procedural memory in internal Merge means that cognition by description places a heavy demand on procedural memory.

But what of the remark one sometimes hears in linguistics that internal Merge 'comes for free'? Does this contradict the point just made? What does it mean to say that internal Merge 'comes for free'? Let's turn to some pertinent literature.

Joseph Aoun and colleagues have argued that the potential use of internal Merge in grammatical derivations is inevitable, given the presence of external Merge and given the distinction between derivations and the lexicon.¹⁷ To quote from them:

We believe that Copy is [...] conceptually necessary, in the sense of following from a very uncontroversial design feature of Universal Grammar. It rests on the fact that there is a (virtually unanimously held) distinction between the lexicon and the computational system and that words are accessed from the lexicon. How does Copy follow from this fact? It is universally assumed that the atoms manipulated by the computational system come from the lexicon. How does the computational system access the lexicon? It does so by copying elements from the lexicon to the computational system. That accessing the lexicon involves copying is clear from the fact that the lexicon gets no smaller when it is accessed and words are obtained for manipulation by the syntax. If this is correct, then grammars that distinguish the lexicon from the computational system conceptually presuppose an operation like Copy. As virtually every approach to grammar assumes something like a distinction between lexicon and grammar, Copy is a 'virtually conceptually necessary' operation [...].
(Aoun *et al.* 2001; cf. Hornstein 2001: 211f.)

Given that "copies are conceptually costless" (Hornstein 2001: 22, n. 10), then the presence of external Merge gives us internal Merge for free. Why? Because internal Merge is Copy combined with (what would otherwise be) external Merge as illustrated earlier by the example of *what Socrates thought what* (Hornstein 2001).

But internal Merge comes for free only as a potential. Internal Merge may exist in the system simply as the existence of Copy and the existence of external Merge. But this alone would be a matter of competence, not execution. In other words, it does not follow that the two would be executed together. The system may not be able to execute internal Merge *in performance* until the procedural

¹⁷ Unless there is a rule forbidding internal Merge, but Ockhamist considerations militate against supposing so.

system has become powerful enough to support a robust derivational memory. Hence, while Aoun *et al.* have made a case for internal Merge in competence, the conclusion of their argument is still compatible with there having been an earlier period of human existence in which external Merge was in use but without internal Merge, due simply to a less developed procedural system.

Merge creates hierarchical structures and hence plausibly relies upon procedural memory, as do all the thought processes which utilize Merge. Internal Merge takes advantage of an especially sophisticated form of procedural memory insofar as it requires memory of derivational steps. Given the importance of quantifier raising in cognition by description, we can speculate that uniquely human procedural memory plays an especially important role in cognition by description, and hence in all the cultural achievements which plausibly depend upon it: awareness of history, religion, and science.

5. The Neuroscience of Procedural Memory

We can make some plausible conjectures about some of the brain mechanisms which underpin cognition by description by considering the neuroscience of procedural memory.

The procedural memory system consists of parallel closed loops between the cortex and the basal ganglia (the corticostriatal circuits), and between the cortex and cerebellum (the corticocerebellar circuits). The corticostriatal circuits consist of parallel and closed loops that project from the cortex to the striatum. Subsequently each circuit splits into two kinds of pathways — direct and indirect — and projects back to the same region of the cortex from which it originated, via the thalamus. The direct pathway projects from the striatum to the globus pallidus interna (GPi), and from there to the substantia nigra and from there to the thalamus. The indirect pathway in turn projects to the globus pallidus externa (GPe), then to the subthalamic nucleus and from there to GPi and then to the thalamus. The different basal ganglia-thalamocortical loops project to different areas of the cortex (e.g., primary motor cortex, premotor cortex and prefrontal cortex) and hence subserve different functions. Indeed, different channels enjoy a similar synaptic organization; this indicates that diverse functions served by the procedural memory system depend on *similar mechanisms*. Each channel is involved in those functions that are carried out by the cortical area to which it projects. The circuits projecting to the primary motor cortex or premotor cortex are involved in motor functions, whereas circuits projecting to the prefrontal cortex are involved in cognitive functions (for review, see Ullman 2004, Ullman & Pierpont 2005).

The cerebellum is also considered very important in procedural memory. The connections between the cerebellum and the cortex are also mostly parallel and functionally segregated. The projections from the cortical areas reach the pontine nuclei; from there the neurons project to the cerebellar cortex. The projections continue into the deep cerebellar nuclei especially the dentate nucleus, and from there to the thalamus and finally again to the cortical area of origin (Kelly & Strick 2003, Middleton & Strick 2000, 2001, Ramnani & Miall 2001,

Ramnani 2006). The corticocerebellar connections are also organized into parallel closed loops. The cerebellum participates in motor learning through sensing and correcting 'motor errors' — i.e. differences between intended movements and those actually performed (Ramnani 2006, Apps & Garwicz 2005). It has been suggested that the cortical regions send copies of their original commands to the cerebellum (called 'efference copies') (for a recent review, see Ramnani 2006).

An important feature of the cerebellum is the uniformity of its cellular organization and circuitry. Therefore, the cerebellum performs *the same computations for every function that it serves*; the reason for the cerebellum being involved in many different functions (including cognitive ones) lies in the different cytoarchitectonic organizations of the cortical areas from which it receives its inputs (Apps & Garwicz 2005, Ramnani 2006).

For a long time it had been supposed that the cerebellum and basal ganglia are involved solely in motor control and that they receive inputs from different areas of the cortex — including the prefrontal cortex which serves for cognitive functions — but send all of their outputs to the motor cortex. However, later findings showed that corticostriatal and corticocerebellar circuits also project to prefrontal cortex and hence may enter into cognitive functions as well (Leiner *et al.* 1993, Dezmund & Fiez 1998, Middleton & Strick 2001, 2002, Gebhart, Petersen & Thach 2002, Kelly & Strick 2003, Ramnani 2006).

In a study conducted by Schoenemann and colleagues, it was shown that while the amount of gray matter in the prefrontal cortex (the area of the cortex serving mainly for cognitive functions) does not differ much between human and nonhuman primates, prefrontal white matter differs greatly (Schoenemann *et al.* 2005). Gray matter is composed of the cell bodies of neurons, whereas white matter is composed of fibers. Moreover, a later study showed that there exists a relatively large prefrontal contribution to the corticocerebellar circuitry in humans when compared to Macaque monkeys. However, in Macaque monkeys the dominant contribution to corticocerebellar circuitry was from the cortical motor areas (Ramnani *et al.* 2006).

These findings suggest that there occurred a selective increase in inter-connectivity between the prefrontal cortex and the basal ganglia and cerebellum (i.e. the procedural memory system) in the human lineage, when compared to non-human primates. Hence it would be quite plausible to suggest that the circuits formerly mainly serving for motor functions (i.e. corticostriatal and corticocerebellar circuits, or stated otherwise 'procedural circuitry') were recruited for cognitive functions in the human lineage. This, in turn, may have played an important role in the great computational power of syntax in language and, as a consequence, to the qualitatively different computational, and ultimately conceptual, powers of the human mind.

6. The Potential for Testing

People suffering from Broca's aphasia, generally understood to be a syntactic disorder, exhibit an interesting lack of abstract thought. In his study of Broca's aphasics, Kurt Goldstein distinguishes two attitudes, the abstract and the

concrete, observing that those with Broca's aphasia tend to be limited to the latter.

In the concrete attitude we are given over passively and bound to the immediate experience of unique objects or situations. Our thinking and acting are determined by the immediate claims made by the particular aspect of the object or situation. For instance, we act concretely when we enter a room in darkness and push the button for light. If, however, we desist from pushing the button, reflecting that by pushing the button we might awaken someone asleep in the room, then we act abstractively. We transcend the immediately given specific aspect of sense impressions, we detach ourselves from the latter and consider the situation from a conceptual point of view and react accordingly. (Goldstein 1948: 6)

Is this lack of abstractness a deficit in cognition by description? Technically no, since we defined cognition by description as the ability to think about entities or agents that one has never seen. We did not define it as the ability to think about states of affairs or situations which one has never perceived. Merge as such, by virtue of being productive, makes possible novel mental representations, so even external Merge, without internal Merge, could perhaps account for the ability to conceive of unperceived situations. The mere presence of a recursive procedure as such may be enough to explain abstractive thought, in Goldstein's sense, as a recursive deficit may also suffice to explain a lack thereof. The hypothesis of this paper, by contrast, predicts that defects in the procedural system will result in difficulties with grammatical transformations as well as difficulties in conceiving of *entities* (and agents) which have never been perceived. The distinction is important to bear in mind while looking for possible evidence.

Piattelli-Palmarini & Uriagereka (2005) conjecture that the uniquely hominid mutation of the *FOXP2* gene, which plausibly led to a boost in procedural memory, made possible transformational grammar (i.e. internal Merge) and hence a wide range of uniquely human cognitive abilities. Our hypothesis is compatible with theirs, although not identical to it. For one thing, we do not put so much weight on *FOXP2*.¹⁸ Perhaps *FOXP2* alone accounts for uniquely hominid, or even uniquely human, procedural memory, but we are also open to roles for other genes as well (Özçelik *et al.* 2008, Tan *et al.* 2008, and references). Furthermore, Piattelli-Palmarini & Uriagereka do not discuss the relevance of Russell to these questions of human evolution.

Our hypothesis is testable. Aphasias have already been discussed. Clear evidence of an aphasia which disables internal Merge, or even just quantifier raising in particular, while leaving cognition by description unimpaired, would refute our hypothesis. 'Clear evidence', however, is an important qualification, because there might be a condition in which internal Merge remains intact but

¹⁸ As Piattelli-Palmarini & Uriagereka (2005: 60) write:

As it turns out, it is not crucial that specifically *FOXP2* be involved in our hypothesis, but since this is the only gene we actually know for sure to be implicated in the language system, *largely for concreteness* we will articulate the proposal around it, and in particular the putative 'permissive' role of *FOXP2* in procedural memory.

cannot be applied in communication. A general inability to think recursively, or even just a general inability to exhibit the computational abilities required for transformational grammar, accompanied by unimpaired cognition by description would offer a clearer refutation.

An interesting potential field of research is to investigate the relation between basal ganglia impairment and deficits in cognition by description. There is some correlation between advanced stages of schizophrenia, basal ganglia dysfunction, and dementia. Since it is late developing, the demented condition is sometimes called 'tardive dementia' (Breggin 1990) or 'tardive dysmentia' (Wilson *et al.* 1983). The condition may be due to schizophrenia being partly a basal ganglia disorder (Graybiel 1997), or it may be a result of anti-schizophrenia medication damaging the basal ganglia (Breggin 1990, 1993, Dalgalarondo & Gattaz 1994), or both. Either way, we submit that it is worthwhile to look for difficulties in the transformational aspects of grammar and impairment in cognition by description in individuals with subcortical dementia.

7. Conclusion

Our conclusion is that Bickerton is mistaken in insisting that uniquely human conceptual structure, specifically cognition by description, must have evolved prior to the evolution of syntax. One can see how the emergence of recursion could have suddenly made possible cognition by description along with syntax. This is consistent with the hypothesis that Merge ushered in both syntax and uniquely human semantics, an hypothesis which Chomsky favors presumably because of its simplicity. However, we also leave open the possibility that external Merge appeared first, meaning that there was a semi-syntax prior to the evolution of full-blown syntax. Specifically, this would have been a phrase-structure grammar without transformations. It would also have been a 'protolanguage' in some sense, but not the a-grammatic sort of protolanguage which Bickerton posited in the earlier quotes. Full-blown syntax, because it utilizes internal Merge, could not have been utilized until a fully developed procedural memory system was in place. So it is possible that the evolution of the memory systems placed a constraint on the evolution of syntax, and hence uniquely human semantics as well, including cognition by description.

Our discussion has been extremely speculative and exploratory, as we noted at the outset. The evidence adduced could, no doubt, be interpreted in other ways. But our aim has been to arrive at a possible explanation of uniquely human semantics, an explanation which can be tested and will as a result of testing, almost certainly, be replaced by something better in time. Our rationale for proposing something so tentative is that one must speculate in order to have something to test. One cannot rule out hypotheses without having hypotheses in the first place. We agree wholeheartedly with Bickerton (2005: 2), when he writes that "Speculation is the horse that drags the chariot of theory".

References

- Anglin Jeremy M. & George A. Miller. 1968. The role of phrase structure in the recall of meaningful verbal material. *Psychonomic Science* 10(10), 343-344.
- Aoun, Joseph, Lina Choueiri & Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32, 371-403.
- Apps, Richard & Martin Garwicz. 2005. Anatomical and physiological foundations of cerebellar information processing. *Nature Reviews Neuroscience* 6, 297-311.
- Bergman, Thore J., Jacinta C. Beehner, Dorothy L. Cheney & Robert M. Seyfarth. 2003. Hierarchical classification by rank and kinship in baboons. *Science* 302, 1234-1236.
- Berwick, Robert C. 1998. Language evolution and the minimalist program: The origins of syntax. In James R. Hurford, Michael Studdert-Kennedy & Chris Knight (eds.), *Approaches to the evolution of language: Social and cognitive bases*, 320-340. Cambridge: Cambridge University Press.
- Bickerton, Derek. 2005. Chomsky: Between a Stony Brook and a hard place. Ms. [http://radicalanthropologygroup.org/pub/bickerton_on_chomsky.pdf (18 September 2008).]
- Bickerton, Derek. 2007. Language evolution: A brief guide for linguists. *Lingua* 117, 510-526.
- Bird, Helen, Matthew A. Lambdon Ralph, Mark S. Seidenberg, James L. McClelland & Karalyn Patterson. 2003. Deficits in phonology and past-tense morphology: What's the connection? *Journal of Memory and Language* 48, 502-526.
- Boolos, George S., John P. Burgess & Richard C. Jeffrey. 2002. *Computability and Logic*, 4th edn. Cambridge: Cambridge University Press.
- Breggin, Peter R. 1990. Brain damage, dementia, and persistent cognitive dysfunction associated with neuroleptic drugs: Evidence, etiology, implications. *Journal of Mind and Behavior* 11, 425-464.
- Breggin, Peter R. 1993. Parallels between neuroleptic effects and lethargic encephalitis: The production of dyskinesias and cognitive disorders. *Brain and Cognition* 23, 8-27.
- Byrne, Richard W. & Anne E. Russon. 1998. Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences* 21, 667-721.
- Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2, 113-124.
- Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2, 137-167.
- Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2(4), 303-351.
- Chomsky, Noam. 1980. *Rules and Representations*. Oxford: Blackwell.
- Chomsky, Noam. 2002. *Syntactic Structures*. Berlin: Mouton de Gruyter. [First published by The Hague: Mouton, 1957.]
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1), 1-22.
- Chomsky, Noam. 2006. Biolinguistic explorations: Design, development,

- evolution. Paper presented at the *Department of Psychology's Distinguished Lecture Series and the Human Development Program Colloquium Series*, Boston University (January 2006). [<http://www.buworldofideas.org/shows/2006/06/20060604.aspS>] (18 September 2008).]
- Chomsky, Noam. 2007. Of minds and language. *Biolinguistics* 1, 9-27.
- Dalgalarrrondo, Paulo. & Wagner F. Gattaz. 1994. Basal ganglia abnormalities in tardive dyskinesia: Possible relationship with duration of neuroleptic treatment. *European Archives of Psychiatry and Clinical Neuroscience* 244, 272-277.
- Devitt, Michael. 2004. The case for referential descriptions. In Marga Reimer & Anne Bezuidenhout (eds.), *Descriptions and Beyond*, 280-305. Oxford: Oxford University Press.
- Donnellan, Keith. 1966. Reference and definite descriptions. *Philosophical Review* 75, 281-304.
- Donnellan, Keith. 1968. Putting humpty dumpty together again. *Philosophical Review* 77, 203-215.
- Epstein, William. 1961a. The influence of syntactical structure on learning. *American Journal of Psychology* 74, 80-85.
- Epstein, William. 1961b. A further study of the influence of syntactical structure on learning. *American Journal of Psychology* 75, 121-126.
- Fitch, W. Tecumseh & Marc D. Hauser. 2004. Computational constraints on syntactic processing in nonhuman primates. *Science* 303, 377-380.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the narrow language faculty: Clarifications and implications. *Cognition* 97, 179-210.
- Fodor, Jerry A. 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fodor, Jerry A. & Thomas G. Bever. 1965. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behaviour* 4, 414-420.
- Gallistel, C. Randy. 1990. Representations in animal cognition: An introduction. *Cognition* 37, 1-22.
- Gebhart, Andrea L., Steven E. Petersen & W. Thomas Thach. 2002. Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Sciences* 978, 318-333.
- Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440(7088), 1204-1207.
- Gibson, Kathleen R. 1993. Tool use, language and social behavior in relationship to information processing capacities. In Kathleen R. Gibson & Tim Ingold (eds.), *Tools, Language and Cognition in Human Evolution*, 251-269. Cambridge: Cambridge University Press.
- Goldstein, Kurt. 1948. *Language and Language Disturbances: Aphasic Symptom Complexes and Their Significance for Medicine and Theory of Language*. New York: Grune & Statton.
- Goudarzi, Sara. 2006. Chomsky challenge: Bird said to grasp key aspect of language. *LiveScience*. [http://www.livescience.com/animals/060426_starling_linguist.html] (18 September 2008).]
- Graf, Richard & Jane W. Torrey. 1966. Perception of phrase structure in written

- language. *Proceedings of the Annual Convention of the American Psychological Association*, 83-84.
- Graybiel, Ann M. 1997. The basal ganglia and cognitive pattern generators. *Schizophrenia Bulletin* 23, 459-469.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569-1579.
- Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Malden, MA & Oxford: Blackwell.
- Hornstein, Norbert & Juan Uriagereka. 1999. Labels and projections: A note on the syntax of quantifiers. *University of Maryland Working Papers in Linguistics* 8, 249-270. [Reprinted in Uriagereka, Juan. 2002. *Derivations: Exploring the Dynamics of Syntax* (Routledge Leading Linguists), 115-135. London: Routledge.]
- Joanisse, Marc F. & Mark S. Seidenberg. 1999. Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America* 96, 7592-7597.
- Johnson, Neal F. 1965. The psychological reality of phrase structure rules. *Journal of Verbal Learning and Verbal Behaviour* 4, 469-475.
- Karmiloff-Smith, Annette. 1992. *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.
- Keenan, Edward L. & Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9, 253-326.
- Kelly Roberta M. & Peter L. Strick. 2003. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *The Journal of Neuroscience* 23(23), 8432-8444.
- Kim, Jaegwon. 1993. *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.
- Larson, Richard & Gabriel Segal. 1998. *Knowledge of Meaning: An Introduction to Semantic Theory*. New Delhi: Prentice-Hall of India. [First published by Cambridge, MA: MIT Press, 1995.]
- Leiner, Henrietta C., Alan L. Leiner & Robert S. Dow. 1993. Cognitive and language functions of the human cerebellum. *Trends in Neurosciences* 16(11), 444-447.
- Levelt, Willem J.M. 1970. Hierarchical chunking in sentence processing. *Perception and Psychophysics* 8, 99-103.
- Longworth, C.E., S.E. Keenan, R.A. Barker, W.D. Marslen-Wilson & L.K. Tyler. 2005. The basal ganglia and rule-governed language use: Evidence from vascular and degenerative conditions. *Brain* 128, 584-596.
- McClelland, James L. & Karalyn Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6, 465-472.
- McGonigle, Brendan, Margaret Chalmers & Anthony Dickinson. 2003. Concurrent disjoint and reciprocal classification by Cebus Apella in seriation tasks: Evidence for hierarchical organization. *Animal Cognition* 6, 185-189.
- Mehler, Jacques, Thomas G. Bever & Peter Carey. 1967. What we look at when we read. *Perception and Psychophysics* 2(5), 213-218.

- Middleton Frank A. & Peter L. Strick. 2000. Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research Reviews* 31(2-3), 236-250.
- Middleton Frank A. & Peter L. Strick. 2001. Cerebellar projections to the prefrontal cortex of the primate. *The Journal of Neuroscience* 21(2), 700-712.
- Middleton Frank A. & Peter L. Strick. 2002. Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cerebral Cortex* 12(9), 926-935.
- Miller, George A. 1962. Some psychological studies of grammar. *American Psychologist* 17, 748-762.
- Miller, George A. S & Stephen Isard. 1963. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behaviour* 2, 217-228.
- Morford, Jill P. & Susan Goldin-Meadow. 1997. From here and now to there and then: The development of displaced reference in homesign and English. *Child Development* 68(3), 420-435.
- Neale, Stephen. 1990. *Descriptions*. Cambridge, MA: MIT Press.
- Newman Aaron J., Michael T. Ullman, Roumyana Pancheva, Diane L. Waligura & Helen J. Neville. 2007. An ERP study of regular and irregular English past tense inflection. *NeuroImage* 34, 435-445.
- Özçelik, Tayfun, Nurten Akarsu, Elif Uz, Şafak Çağlayan, Süleyman Gülsüner, Onur Emre Onat, Meliha Tan & Üner Tan. 2008. Mutations in the very low-density lipoprotein receptor *VLDLR* cause cerebellar hypoplasia and quadrupedal locomotion in humans. *Proceedings of the National Academy of Sciences of the United States of America* 105, 4232-4236.
- Parker, Anna R. 2006. *Evolution as a Constraint on Theories of Syntax: The Case against Minimalism*. Edinburgh: University of Edinburgh dissertation.
- Piattelli-Palmarini, Massimo & Juan Uriagereka. 2005. The evolution of the narrow faculty of language: The skeptical view and a reasonable conjecture. *Lingue e Linguaggio* IV, 27-79.
- Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: What's so special about it? *Cognition* 95(2), 201-236.
- Pinker, Steven & Michael T. Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6(11), 456-463.
- Ramnani, Narender. 2006. The primate cortico-cerebellar system: Anatomy and function. *Nature Reviews Neuroscience* 7(7), 511-522.
- Ramnani, Narender, Timothy E.J. Behrens, Heidi Johansen-Berg, Marlene C. Richter, Mark A. Pinsk, Jesper L.R. Andersson, Peter Rudebeck, Olga Ciccarelli, Wolfgang Richter, Alan J. Thompson, Charles G. Gross, Matthew D. Robson, Sabine Kastner & Paul M. Matthews. 2006. The evolution of prefrontal inputs to the cortico-pontine system: Diffusion imaging evidence from Macaque monkeys and humans. *Cerebral Cortex* 16(6), 811-818.
- Ramnani Narender & Chris Miall. 2001. Expanding cerebellar horizons. *Trends in Cognitive Sciences* 5(4), 135-136.
- Reinhart, Tanya. 1976. *The Syntactic Domain of Anaphora*. Cambridge, MA: MIT dissertation.
- Russell, Bertrand. 1905. On denoting. *Mind* 14, 479-493.
- Russell, Bertrand. 1910. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society* 11, 108-128.
- Russell, Bertrand. 1919. *Introduction to Mathematical Philosophy*. London: Allen and

Unwin.

- Russell, Bertrand. 1957. Mr. Strawson on referring. *Mind* 66, 385-389.
- Russell, Bertrand. 1959. *The Problems of Philosophy*. Oxford: Oxford University Press.
- Schoenemann P. Thomas, Michael J. Sheehan & L. Daniel Glotzer. 2005. Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nature Neuroscience* 8(2), 242-252.
- Smith, Neil. 2004. *Chomsky: Ideas and Ideals*, 2nd edn. Cambridge: Cambridge University Press.
- Spinozzi, Giovanna, Francesco Natale, Jonas Langer & Karen E. Brakke. 1999. Spontaneous class grouping behavior by bonobos (*Pan Paniscus*) and common chimpanzees (*P. Troglodytes*). *Animal Cognition* 2, 157-170.
- Squire, Larry R. 2004. Memory systems of the brain: A brief history and current Perspective. *Neurobiology of Learning and Memory* 82, 171-177.
- Stromswold, Karin. 1999. Cognitive and neural aspects of language acquisition. In Ernest Lepore & Zenon Pylyshyn (eds.) *What is cognitive science?*, 356-400). Oxford: Blackwell.
- Suzuki, Ryuji, John R. Buck & Peter L. Tyack. 2006. Information entropy of humpback whale song. *The Journal of the Acoustical Society of America* 119(3), 1849-1866.
- Tan, Üner, Sibel Karaca, Meliha Tan, Bekir Yılmaz, Namık Kemal Bağcı, Ayhan Özkur, & Sadrettin Pence. 2008. Unertan syndrome: A case series demonstrating human devolution. *International Journal of Neuroscience* 118, 1-25.
- Thach, W. Thomas & Amy J. Bastian. 2004. Role of the cerebellum in the control and adaptation of gait in health and disease. *Progress in Brain Research* 143, 353-366.
- Turing, Alan M. 1950. Computing machinery and intelligence. *Mind* 59, 433-460.
- Ullman, Michael T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92, 231-270.
- Ullman, Michael T. & Myrna Gopnik. 1994. Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics* 20, 51-117.
- Ullman, Michael T. & Elizabeth I. Pierpont. 2005. Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex* 41(3), 399-433.
- Watkins, Kate E., Faraneh Vargha-Khadem, John Ashburner, Richard E. Passingham, Alan Connelly, Karl J. Friston, Richard S.J. Frackowiak, Mortimer Mishkin & David G. Gadian. 2002. MRI analysis of an inherited speech and language disorder: Structural brain abnormalities. *Brain* 125, 465-478.
- Wilson, Ian C., James C. Garbutt, Clayton F. Lanier, Joseph Moylan, William Nelson & Arthur J. Prange. 1983. Is there a tardive dysmentia? *Schizophrenia Bulletin* 9, 187-192.

*John Bolender**Middle East Technical University
Department of Philosophy
Eskişehir Yolu
06531 Ankara
Turkey*bolender@metu.edu.tr*Burak Erdeniz**Informatics Institute
İnönü Bulvarı
06531 Ankara
Turkey*burakerdeniz@yahoo.com*Cemil Kerimoğlu**Georg-August-Universität Göttingen
International Max Planck Research School
Neuroscience Program, Grisebachstr. 5
37077 Göttingen
Germany*gencemil@gmail.com